# Structure of the $n$th Roots of a Matrix

Gabriëlle ten Have
*Mathematical Institute*
*University of Leiden*
*P. O. Box 9512*
*2300 RA Leiden, the Netherlands*

Submitted by Thomas J. Laffey

ABSTRACT

Let $K$ be a subfield of $\mathbf{C}$. We give a criterion for a nonsingular matrix $A$ in $M_m K$ to have an $n$th root in $M_m K$. The number of similarity classes of $n$th roots of $A$ is given for $K = \mathbf{R}$. Further we indicate which matrices $A$ in $M_m K$ have infinitely many $n$th roots in $M_m K$.

## 0. INTRODUCTION

Let $A$ be a nonsingular matrix in $M_m \mathbf{Z}$ or in $M_m K$, $m \geqslant 2$, where $K$ is an arbitrary subfield of $\mathbf{C}$. We are interested in the number of solutions or solution classes in $M_m \mathbf{Z}$ and $M_m K$ of the matrix equation $X^n = A$. Recently, Otero [6] published an article about the extraction of roots in matrix rings over an arbitrary field $F$. He gives a module-theoretic criterion for a matrix in $M_m F$ to have an $n$th root. In this paper we give a more constructive way of finding roots in $M_m K$. Further we answer a question of Otero by giving the number of similarity classes of real $n$th roots of a real matrix. It is also shown that, assuming the existence of an $n$th root $X$ in $M_m K$ of a matrix $A$ in $M_m K$, the existence of infinitely many complex $n$th roots similar to $X$ of $A$ implies the existence of infinitely many $n$th roots of $A$ in $M_m K$.

Results on the more general matrix equation $f(X) = A$ where $A \in M_m \mathbf{C}$ and $f$ is a complex holomorphic function defined on an open subset of $\mathbf{C}$ can be found in [2] and in a recently published article of Evard and Uhlig [1]. In the latter paper a wide range of results on the structure of complex and real




0024-3795/93/$6.00

solutions is given. One of the results, mentioned in [1, Corollaries 4.7 and 4.8], includes the corollary stated in Section 2.

## 1. THE COMPLEX CASE

When $K = \mathbf{C}$ the structure of the solutions of $X$ of $X^n = A$ is known. Let the Jordan canonical form of $A$ be given by

$$ZAZ^{-1} = J_A = J_1 \oplus J_2 \oplus \cdots \oplus J_s = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s \end{pmatrix}.$$

Let $\alpha_i$ be the eigenvalue corresponding to the $m_i \times m_i$ Jordan block $J_i$, and $Z \in M_m\mathbf{C}$.

For $x \in \mathbf{C}$ let $g(x) = x^{1/n}$ denote any value of the $n$th root of $x$, and let $g(J_i)$ denote any of the $n$th roots of $J_i$ (see for instance [2, p. 232]).

When $A$ is non-singular, all solutions $X$ of $X^n = A$ are given by the matrices

$$\begin{aligned} g(A) &:= Z^{-1}U^{-1}g(J_A)UZ \\ &= Z^{-1}U^{-1}\begin{pmatrix} g(J_1) & & & \\ & g(J_2) & & \\ & & \ddots & \\ & & & g(J_s) \end{pmatrix}UZ \end{aligned}$$

for all nonsingular $U \in M_m\mathbf{C}$ satisfying $UJ_A = J_AU$ [2, p. 232]. When all eigenvalues $\alpha_i$ of $A$ are distinct, the matrices commuting with $J_A$ are polynomials in $J_A$ [5, p. 419, Proposition 1] and thus polynomials in $g(J_A)$. These matrices commuting with $J_A$ must consequently commute with $g(J_A)$, too. Therefore we find exactly $n^s$ solutions of $X^n = A$. These solutions can be written as polynomials in $A$ [2, p. 233].

According to Lancaster and Tismenetsky [5, p. 418, Theorem 1], when not all of the eigenvalues are distinct, there is at least one $g(J_A)$ for which an infinity of nonsingular matrices $U$ exists such that $U$ commutes with $J_A$ but not with the Jordan canonical form $J_g$ of $g(J_A)$. Since $U$ is a block matrix, partitioned as $J_g$ and $g(J_A)$ and consisting of upper-triangular Toeplitz matrices [5, p. 419], $U$ commutes with $g(J_A)$ if and only if it commutes with $J_g$. Thus $U$ does not commute with $g(J_A)$ either. In Theorem 2 we will prove

that these $U$ induce infinitely many distinct solutions $X$ of $X^n = A$. Each $g(J_A)$ represents one similarity class over $\mathbf{C}$, whence we find exactly $n^s$ similarity classes over $\mathbf{C}$, some of which are finite.

## 2. ROOTS IN A FIELD $K$

Consider a subfield $K$ of $\mathbf{C}$. We will first introduce some terminology.

For any matrix $A \in M_m K$ and any $k \in \{1, \ldots, m\}$, define $D_k(\lambda)$ as the g.c.d. in $K[\lambda]$ of the determinants of all $k \times k$ submatrices of $\lambda I - A$. Set $D_0(\lambda) \equiv 1$. The quotients $i_k(\lambda) = D_k(\lambda)/D_{k-1}(\lambda)$, which are polynomials in $\lambda$, are called the invariant polynomials of $A$. Let for each $k$ the factorization over $K$ of the polynomials $i_k(\lambda)$ in irreducible factors be given by $i_k(\lambda) = f_{k,1}(\lambda)^{k_1} \cdots f_{k,j_k}(\lambda)^{k_{j_k}}$. Then we call the polynomials $f_{k,i}(\lambda)^{k_i}$ with $k = 1, \ldots, m$ and $i = 1, \ldots, j_k$ the elementary divisors of $A$ in $K[\lambda]$. The Jordan blocks of a matrix $A \in M_m\mathbf{C}$ can be found directly from the elementary divisors of $A$ [2, p. 151]. Further, two matrices $A, B \in M_m K$ are similar if and only if $A$ and $B$ have the same invariant polynomials [2, p. 197, Theorem 10].

When $h(x) = x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0 \in K[x]$, the companion matrix $C(h)$ is defined as the $m \times m$ matrix

$$C(h) = \begin{pmatrix} -a_{m-1} & -a_{m-2} & \cdots & -a_1 & -a_0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix},$$

which has $h(x)$ as its characteristic polynomial and as its minimum (annihilating) polynomial. Suppose $h(x)$ is irreducible over $K$ with factorization over $\mathbf{C}$ given by $h(x) = (x - \alpha_1) \cdots (x - \alpha_m)$. By definition we have, for any positive integer $k$, that $h^k$ is the minimum polynomial of $C(h^k)$. It follows that $C(h^k)$ is similar to the Jordan canonical form $J(\alpha_1) \oplus \cdots \oplus J(\alpha_m)$, where each $J(\alpha_i)$ is a $k \times k$ Jordan block.

Let $\oplus dA$ stand for the block-diagonal matrix $A \oplus A \oplus \cdots \oplus A$ where $A$ appears $d$ times. Then a matrix $A \in M_m K$ has a $K$-normal form $C_A = \oplus_{i=1}^t d_i C(f_i^{k_i})$, where the $f_i^{k_i}(x)$ are the elementary divisors in $K[x]$ of $A$ [2, pp. 149–150]. By the definition of elementary divisor the $f_i(x)$ are irreducible, and by using the aforementioned multiplicity $d_i$ we can assume that no two of the $f_i(x)^{k_i}$ are the same. Note that, when $m_i$ is the degree of $f_i(x)$, the order $m$ of the matrix $A$ is equal to $\Sigma_{i=1}^t d_i k_i m_i$. Now, for every $i$, take an $\alpha_i = \alpha_i^{(1)}$ which satisfies $f_i(\alpha_i) = 0$, and for $j = 1, \ldots, n$, take the $n$ distinct

values $\gamma_{ij}$ such that $\gamma_{ij}^n = \alpha_i$. Let $n_{ij}$ be given by $n_{ij} = [K(\gamma_{ij}) : K(\alpha_i)]$ for $i = 1, \ldots, t$ and $j = 1, \ldots, n$.

THEOREM 1. *Let $K$ be a subfield of $\mathbf{C}$, and let $A \in M_m K$ be nonsingular. Then, with notation as above, $A$ has an $n$th root in $M_m K$ if and only if there are nonnegative integers $b_{i1}, \ldots, b_{in}$ such that $d_i = b_{i1}n_{i1} + b_{i2}n_{i2} + \cdots + b_{in}n_{in}$ for $i = 1, \ldots, t$.*

*Proof.* Necessity: When for $b_{i1}, \ldots, b_{in} \in \mathbf{Z}_{\geq 0}$ a block $C(f_i^{k_i})$ appears $b_{i1}n_{i1} + \cdots + b_{in}n_{in}$ times in $C_A$, this means that for every root $\alpha_i^{(l)}$ of $f_i(x) = 0$ $(l = 1, \ldots, m_i)$ the irreducible Jordan block $J(\alpha_i^{(l)})$ appears $b_{i1}n_{i1} + \cdots + b_{in}n_{in}$ times in the Jordan canonical form of $A$. Since $f_i(x)$ is irreducible, all $K(\alpha_i^{(l)})$ are isomorphic to $K(\alpha_i)$. We are going to choose $n$th roots $\gamma_{i1}^{(1)}, \ldots, \gamma_{id_i}^{(1)}, \ldots, \gamma_{i1}^{(m_i)}, \ldots, \gamma_{id_i}^{(m_i)}$ of $\alpha_i^{(1)}, \ldots, \alpha_i^{(m_i)}$, respectively, such that these $\gamma_{ij}^{(l)}$ satisfy a polynomial of degree $m_i d_i$ in $K[x]$. This is done as follows. For $j = 1, \ldots, n$ we take $\gamma_{ij}^{(1)}$ and each of the other $n_{ij} - 1$ roots having the same minimum polynomial over $K(\alpha_i^{(1)})$ as $\gamma_{ij}^{(1)}$. Now for each of the other $m_i - 1$ roots of $f_i(x) = 0$ we choose the corresponding conjugated roots of the $\gamma_{ij}^{(1)}$. Doing so, we have selected $n_{ij}$ $n$th roots of $\alpha_i^{(1)}, \ldots, \alpha_i^{(m_i)}$. By construction these $m_i n_{ij}$ roots satisfy a polynomial $g_{ij}(x) \in K[x]$, the minimum polynomial over $K$ of $\gamma_{ij}^{(1)}$. Now define $C_{P_i}$ as the block-diagonal matrix $\oplus_{j=1}^n b_{ij} C(g_{ij}^{k_i})$, and $C_P$ as $C_{P_1} \oplus \cdots \oplus C_{P_t} \in M_m K$. To each characteristic root $\alpha_i^{(l)}$ of $A$ corresponds a root $\gamma_{ij}^{(l)}$ with $(\gamma_{ij}^{(l)})^n = \alpha_i^{(l)}$ of $C_P$, and the Jordan blocks of $\alpha_i^{(l)}$ and $\gamma_{ij}^{(l)}$ have the same order $k_i$. Thus $C_P^n$ is similar to $A$, i.e., $BC_P^n B^{-1} = A$, or equivalently $BC_P^n = AB$, for some $B \in M_m \mathbf{C}$. Since $A$, $C_P^n \in M_m K$, the entries of $B$ can be found as solutions of linear equations with coefficients in $K$. Therefore we can find $B \in M_m K$ such that $(BC_P B^{-1})^n = BC_P^n B^{-1} = A$. Take $P = BC_P B^{-1}$.

Sufficiency: Let $P \in M_m K$ be a root of $X^n = A$. The $K$-normal form of $A$ can be written as $C_A = \oplus_{i=1}^t d_i C(f_i^{k_i})$. For each $i$ take a root $\alpha_i$ of $f_i(x) = 0$. These $\alpha_i$ appear in $d_i$ Jordan blocks of $A$, each of size $k_i \times k_i$. Take all $\gamma_{ij}$ satisfying $\gamma_{ij}^n = \alpha_i$. Now each of the $d_i$ $k_i \times k_i$ Jordan blocks corresponding to $\alpha_i$ originates from a $k_i \times k_i$ Jordan block of $P$ corresponding to one of the roots $\gamma_{ij}$. Suppose $P$ has $c_{ij}$ $k_i \times k_i$ Jordan blocks corresponding to the characteristic root $\gamma_{ij}$, where $c_{ij} \geq 0$. We have $\sum_{j=1}^n c_{ij}$ equal to $d_i$, since the characteristic roots of $A$ are the $n$th powers of the characteristic roots of $P$. On the other hand $P \in M_m K$. Hence, by the definition of $n_{ij}$, each $\gamma_{ij}$ must appear in as many $k_i \times k_i$ Jordan blocks of $P$ as its $n_{ij} - 1$ conjugated roots over $K(\alpha_i)$. Thus, taking together blocks that belong to conjugated roots, $d_i = \sum_{j=1}^n e_{ij} c_{ij} n_{ij}$ for some nonnegative integers $e_{ij}$. ∎

For the construction, find the Jordan canonical form of an $n$th root $P$ of $A$ as indicated in the proof. From this form a $K$-normal form $C_P$ can be deduced. Since $C_P^n$ is similar to $A$, a matrix $Y \in M_m K$ exists such that $(Y^{-1}C_P Y)^n = Y^{-1}C_P^n Y = A$. This $Y$ can be found from the equation $C_P^n Y = YA$.

COROLLARY. *Let $m \geqslant 2$. Let $A \in M_m\mathbf{R}$ be nonsingular. Then for odd $n$ a real solution to $X^n = A$ always exists, for even $n$ a real solution exists if and only if each elementary divisor (Jordan block) of $A$ corresponding to a negative eigenvalue occurs an even number of times.*

In order to count the similarity classes of real solutions $X$ of $X^n = A$, let the elementary divisors over $\mathbf{C}$ of $A$ be as follows:

$(\lambda - \lambda_i)^{a_i}$ appears $s_i$ times for $i = 1, \ldots, k$, $\lambda_i \in \mathbf{R}_{>0}$;
$(\lambda - \mu_i)^{b_i}$ appears $t_i$ times for $i = 1, \ldots, l$, $\mu_i \in \mathbf{R}_{<0}$;
$(\lambda - \nu_i)^{c_i}$ and $(\lambda - \bar{\nu}_i)^{c_i}$ appear $u_i$ times for $i = 1, \ldots, m$, $\nu_i \in \mathbf{C} \setminus \mathbf{R}$.

Further, let

$$\delta_n = \begin{cases} 0 & \text{for } n \text{ odd}, \\ 1 & \text{for } n \text{ even} \end{cases}$$

and

$$S_i = \begin{cases} s_i + 1 & \text{for } n = 2 \\ \displaystyle\sum_{j=0}^{[s_i/2]} \binom{\left[\frac{n-1}{2}\right] + j - 1}{j} (s_i - 2j + 1)^{\delta_n} & \text{for } n \geqslant 3, \end{cases}$$

$$T_i = \begin{cases} 0 & \text{for } n \text{ even}, \ t_i \text{ odd} \\ \displaystyle\binom{\frac{n}{2} + \frac{t_i}{2} - 1}{\frac{t_i}{2}} & \text{for } n \text{ even}, \ t_i \text{ even} \\ \displaystyle\sum_{j=0}^{[t_i/2]} \binom{\frac{n-1}{2} + j - 1}{j} & \text{for } n \geqslant 3 \text{ odd}, \end{cases}$$

$$U_i = \binom{n + u_i - 1}{u_i}.$$

PROPOSITION. *With notation as above, for $n \geqslant 2$ the number of similarity classes of real nth roots of the real matrix A is given by*

$$\left(\prod_{i=1}^{k} S_i\right) \cdot \left(\prod_{i=1}^{l} T_i\right) \cdot \left(\prod_{i=1}^{m} U_i\right).$$

*Proof.* Any choice of eigenvalues for a real $n$th root $X$ of $A$ induces a similarity class of real $n$th roots of $A$. Thus we must count the sets of eigenvalues admissible for the real matrix $X$. Now for a characteristic root $\lambda$ of $A$ count the real $n$th roots and the pairs of complex conjugated $n$th roots. Then use the fact that $\binom{d+j-1}{j}$ is the number of $j$-combinations with repetition of $d$ distinct objects [7, p. 7]. ■

## 3. INFINITELY MANY SOLUTIONS

In the previous section we found when a solution $X$ exists of $X^n = A$ with $X, A \in M_m K$. Now we want to know for which $A$ there are infinitely many solutions.

Let $X, A \in M_m K$. When all eigenvalues of $A$ are distinct, there are only finitely many solutions $X$, as we showed in Section 1 for $K = C$. When not all of the eigenvalues are distinct, we can find a $g(J_A)$ and infinitely many nonsingular matrices $U \in M_m\mathbf{C}$ such that $U$ commutes with $J_A$ but not with $g(J_A)$. Since $U, J_A, g(J_A) \in M_m\mathbf{C}$, it is not immediately clear that these matrices $U$ induce infinitely many distinct solutions $X = g(A) \in M_m K$. By construction we will prove that infinitely many such matrices $X$ indeed exist.

As we saw in Section 1, each class of solutions $X$ of $X^n = A$ is determined by one of the $n^s$ possible Jordan canonical forms of $g(J_A)$. Fix a solution $P \in M_m K$ with Jordan canonical form $J_P$. Then all solutions in the same class as $P$ are given by $V^{-1}PV$ where $V$ commutes with $A$.

THEOREM 2. *Let A and P be as above. Suppose there exists a nonsingular matrix $U \in M_m\mathbf{C}$ that commutes with $J_A$ but not with $J_P$. Then there is an infinite set of matrices $\{V_1, V_2, V_3, \ldots\}$ in $M_m K$ with every $V_i$ commuting with A such that the matrices $V_i^{-1}PV_i$ are distinct.*

*Proof.* Let the Jordan canonical forms of $A$ and $P$ be given by $J_A = J(\alpha_1) \oplus \cdots \oplus J(\alpha_s)$ and $J_P = J(\beta_1) \oplus \cdots \oplus J(\beta_s)$ respectively. The $J(\alpha_i)$ and $J(\beta_i)$ are $m_i \times m_i$ Jordan blocks for some $m_i \in \mathbf{Z}$. For $1 \leqslant i, j \leqslant s$ let $a_{ij}$ be the degree of the g.c.d. of the elementary divisors $(x - \alpha_i)^{m_i}$ and

$(x - \alpha_j)^{m_j}$ of $A$, and $b_{ij}$ the degree of the g.c.d. of the elementary divisors $(x - \beta_i)^{m_i}$ and $(x - \beta_j)^{m_j}$ of $P$. According to Lancaster and Tismenetsky [5, p. 418] the linear space of solutions $X$ of $AX = XA$ (and also of $J_A X = X J_A$) has dimension $a = \sum_{i,j=1}^{s} a_{ij}$, and the linear space of solutions $X$ of $PX = XP$ has dimension $b = \sum_{i,j=1}^{s} b_{ij}$. Since we assumed the existence of a matrix $U \in M_m\mathbf{C}$ commuting with $J_A$ but not with $J_P$, we must have $a - b \geqslant 1$. Thus a matrix $V \in M_m\mathbf{C}$ exists commuting with $A$ but not with $P$. As solution of linear equations with coefficients in $K$, $V$ can be taken in $M_m K$. We are going to define a set of matrices $\mathbf{V} = \{V_{\lambda_1}, V_{\lambda_2}, V_{\lambda_3}, \dots\}$ such that for $V \in \mathbf{V}$

$$\begin{cases} AV = VA, \\ PV \neq VP, \\ \det V \neq 0, \\ \text{all } V^{-1}PV \text{ are distinct.} \end{cases} \tag{1}$$

Take $V_\lambda = V + \lambda I$ for $\lambda \in K$. There exists a certain constant $\lambda_0 \in K$ depending only on $V$ such that $V_\lambda$ is nonsingular for $|\lambda| \geqslant \lambda_0$; thus we take $|\lambda| \geqslant \lambda_0$ in the rest of the proof. Clearly $V_\lambda$ fulfills the first three conditions of (1) for all $\lambda$ if $V$ satisfies these conditions. Now assume that $V_{\lambda_1}, \dots, V_{\lambda_j}$ satisfy (1) for certain $j \geqslant 1$. We shall show that we can find a matrix $V_{\lambda_{j+1}}$ such that $V_{\lambda_1}, \dots, V_{\lambda_{j+1}}$ satisfy (1). Note that $V^{-1}PV = W^{-1}PW$ exactly when $(VW^{-1})^{-1}P(VW^{-1}) = P$, i.e., when $VW^{-1}$ commutes with $P$. Thus consider $V_{\lambda_{j+1}}V_{\lambda_k}^{-1} = (V + \lambda_{j+1}I)V_{\lambda_k}^{-1}$ for $\lambda_{j+1} \notin \{\lambda_1, \dots, \lambda_j\}$ and for $k = 1, \dots, j$. Suppose there is a $k$ such that $V_{\lambda_{j+1}}V_{\lambda_k}^{-1}$ commutes with $P$. We know that $\lambda' V_{\lambda_k}^{-1}$ does not commute with $P$ for any $\lambda' \neq 0$. Hence $V_{\lambda_{j+1}+\lambda'}V_{\lambda_k}^{-1} = (V + \lambda_{j+1}I)V_{\lambda_k}^{-1} + \lambda' V_{\lambda_k}^{-1}$ does not commute with $P$ for $\lambda' \neq 0$. Thus, we can choose $V_{\lambda_{j+1}}$ such that $V_{\lambda_1}, \dots, V_{\lambda_{j+1}}$ satisfy (1), since only finitely many values $\lambda \geqslant \lambda_0$ for $\lambda_{j+1}$ are excluded. ∎

## 4. INTEGRAL SOLUTIONS

For a nonsingular matrix $A \in M_m\mathbf{Z}$, several problems arise in the search for integral $n$th roots. Firstly, no existence theorems are known for $m > 2$. Secondly, no general results are known about the number of similarity classes of matrices. For instance, the matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}$$

are similar over $\mathbf{Q}$, but they are not similar over $\mathbf{Z}$.

For $m = 2$, [3] gives all solutions in case $A \neq aI$ for $a \in \mathbf{Z} \setminus \{0\}$, and (an upper bound for) the number of similarity classes in case $A = aI$.

For $m \geqslant 2$, Latimer and MacDuffee [4] connect the number of similarity classes over $\mathbf{Z}$ of the set of nonsingular matrices in $M_m\mathbf{Z}$ having $m$ distinct characteristic roots with the number of ideal classes in a certain order. This result can very well be used in case a matrix root $P$ of $A$ has $m$ distinct characteristic roots, although even then in some cases only a lower bound of the number of similarity classes can be given. In case $P$ has less than $m$ distinct characteristic roots, either because some Jordan blocks have order greater than one or because some Jordan blocks occur more than once, no results are known.

REFERENCES

1 J.-C. Evard and F. Uhlig, on the matrix equation $f(X) = A$, *Linear Algebra Appl.* 162–164:447–519 (1992).
2 F. R. Gantmacher, *Matrix Theory*, Vol. 1, Chelsea, 1959.
3 G. N. ten Have, Matrix solutions of the equation $X^n = A$, *Indag. Math.* (*N.S.*) 2:57–64 (1991).
4 C. G. Latimer and C. C. MacDuffee, A correspondence between classes of ideals and classes of matrices, *Ann. Math.* 34:313–316 (1933).
5 P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, Academic, 1985.
6 D. E. Otero, Extraction of $m$th roots in matrix rings over fields, *Linear Algebra Appl.* 128:1–26 (1990).
7 J. Riordan, *An Introduction to Combinatorial Analysis*, Wiley, 1958.

*Received 12 December 1991; final manuscript accepted 10 August 1992*